Numer. Math. (2003)
Digital Object Identifier (DOI) 10.1007/s00211-003-0482-9

Numerische Mathematik

A posteriori error estimation for finite element discretization of parameter identification problems

Roland Becker, Boris Vexler*

Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294/293 69120 Heidelberg, Germany; e-mail: {roland.becker, boris.vexler}@iwr.uni-heidelberg.de

Received May 27, 2002 / Revised version received January 6, 2003 /
Published online: September 5, 2003 –

Summary. In this paper we develop an a posteriori error estimator for parameter identification problems. The state equation is given by a partial differential equation involving a finite number of unknown parameters. The presented error estimator aims to control the error in the parameters due to discretization by finite elements. For this, we consider the general setting of a partial differential equation written in weak form with abstract parameter dependence. Exploiting the special structure of the parameter identification problem, allows us to derive an error estimator which is cheap in comparison to the overall optimization algorithm. Several examples illustrating the behavior of an adaptive mesh refinement algorithm based on our error estimator are discussed in the numerical section. For the problems considered here, both, the efficiency of the estimator and the quality of the generated meshes are satisfactory.

Mathematics Subject Classifications (2000): 65K10, 65N30, 49K20

1 Introduction

We consider parameter identification problems involving a finite number of unknown parameters in the following abstract form: The state variable u in an appropriate Hilbert space V is determined by a partial differential equation

* This work has been supported by the German Research Foundation (DFG) through SFB 359 Reactive Flow, Diffusion and Transport and Graduiertenkolleg "Modellierung und wissenschaftliches Rechnen in Mathematik und Naturwissenschaften

Correspondence to: R. Becker

(*state equation*) in weak form:

$$a(u,q)(\phi) = f(\phi) \quad \forall \phi \in V, \tag{1}$$

where $q \in Q = \mathbb{R}^{n_p}$ denotes the unknown parameters. The semi-linear form $a(\cdot,\cdot)(\cdot)$ is defined on the Hilbert space $V \times Q \times V$. Semi-linear forms are written with two parentheses, the first one refers to the nonlinear arguments, whereas the second one embraces all linear arguments. The partial derivatives of the semi-linear form $a(\cdot,\cdot)(\cdot)$ are denoted by $a'_u(\cdot,\cdot)(\cdot,\cdot)$, $a'_q(\cdot,\cdot)(\cdot,\cdot)$ etc.

Further, we are given an observation operator $C : V \to Z$, which maps the *state variable* u to the space of measurements $Z = \mathbb{R}^{n_m}$, where we assume that $n_m \geq n_p$. We denote by $\langle\cdot,\cdot\rangle_Z$ the scalar product of Z and by $\|\cdot\|_Z$ the corresponding norm. Similar notation are used for the scalar product and norm in the space Q.

The values of the parameters are estimated from a given set of measurements $\bar{C} \in Z$ using a least squares approach, such that we obtain the constrained optimization problem with the cost functional $J : V \times Q \to \mathbb{R}$:

$$\text{Minimize} \quad J(u,q) := \frac{1}{2}\|C(u) - \bar{C}\|_Z^2 + \frac{\alpha}{2}\|q - \bar{q}\|_Q^2, \tag{2}$$

under the constraint (1). Here, the cost functional (2) is the sum of the squared norm of the so called *least squares residual* defined by

$$R^{LS}(u) := \bar{C} - C(u), \tag{3}$$

and a regularization term involving prescribed $\alpha \geq 0$ and $\bar{q} \in Q$.

The state equation is discretized by conforming finite elements on a regular mesh $\mathcal{T}_h$, resulting in a finite element space $V_h \subset V$ (for precise definitions see Section 2.1). The corresponding discrete state $u_h \in V_h$ and parameter $q_h \in Q$ are determined by:

$$\text{Minimize} \quad J(u_h, q_h) \tag{4}$$

under the constraint

$$a(u_h, q_h)(\phi_h) = f(\phi_h) \quad \forall \phi_h \in V_h. \tag{5}$$

Due to the finite dimension of Q, the parameter q_h in (4) comes from the same space Q.

The goal of this paper is the development of an a posteriori error estimator for the error in the parameters. Its purpose is to guide an adaptive mesh refinement algorithm. Furthermore, the estimator is used to assess the accuracy of the computed parameters. In order to measure the error in the parameters, we introduce an error functional $E : Q \to \mathbb{R}$. The use of the error functional E

allows one to weight the relative importance of the different parameters. We prove the following error representation:

$$E(q) - E(q_h) = \eta_h + P + R, \tag{6}$$

where η_h denotes the a posteriori error estimator and P and R are remainder terms, which may usually be neglected; see the discussion in Section 3. Further, for its use in a mesh adaptation procedure, it is important that the estimator η_h is the sum of cell-wise contributions.

Our error estimator is based on the optimal control approach to a posteriori error estimation developed in Becker & Rannacher [3,4]. However, direct application of this approach leads to an estimator which controls the error in the cost functional (2). In general, such an estimator does not provide useful error bounds for the parameters, in contrast to the approach presented in this article, see (6). The main idea in deriving our error representation is the formulation of a special auxiliary equation, which is presented in Section 3.

In order to illustrate the typical use of the error estimator η_h, we sketch a generic adaptive mesh refinement algorithm. Such an algorithm generates a sequence of locally refined meshes and corresponding finite element spaces until the estimated error with respect to E is below a given tolerance TOL. For the following iteration, we suppose to have a mesh refinement procedure that adaptively refines a given regular mesh to obtain a new regular mesh for the next iteration. The refinement procedure is guided by information based on the cell-wise contributions of the estimator η_h.

Adaptive Mesh Refinement Algorithm

1. Choose an initial mesh $\mathcal{T}_{h_0}$ and set $k = 0$
2. Construct the finite element space V_{h_k}
3. Compute $u_{h_k} \in V_{h_k}$, $q_{h_k} \in Q$ solving (4,5)
4. Evaluate the a posteriori error estimator η_{h_k}
5. If $\eta_{h_k} \leq TOL$ quit
6. Refine $\mathcal{T}_{h_k} \to \mathcal{T}_{h_{k+1}}$ using information from η_{h_k}
7. Increment k and go to 2.

Remark 1 In step 3, the least squares problem is solved on a fixed mesh. As initial data, we use the values from the computation on the previous mesh. This allows us to avoid unnecessary iterations of the optimization loop on fine meshes.

Although the concepts of adaptivity and a posteriori error estimation are now commonly accepted for the numerical solution of partial differential equations, to our knowledge, there are only few published results on adaptive finite elements for optimization problems and in particular for parameter identification problems, see [1,5,13]. In our earlier work [6] we developed the first ideas of our strategy for mesh adaptation for parameter identification problems. There, the solution of certain dual problems in a Gauß-Newton iteration was essential. Here, we present a general approach which, amongst other things, allows us to avoid assumptions on the smallness of the least squares residual $R^{LS}(u)$.

The outline of this article is as follows: In Section 2 we define the finite element discretization of the parameter identification problem on locally refined meshes. We also describe a typical optimization loop which is carried out on a fixed mesh. Section 3 is devoted to the development of the error estimator. First, we discuss a simple linear example in order to illustrate our approach. Next we derive the error estimator in the general nonlinear case without assuming a perfect match, i.e without assuming $C(u) = \bar{C}$. Numerical examples are discussed in Section 4. We consider linear and nonlinear elliptic equations where the parameters enter in different ways.

2 Solution algorithm on a fixed mesh

In this section, we describe a solution algorithm for the general parameter identification problem on a fixed mesh. This includes reformulation of the problem (1,2) as an unconstrained optimization problem, discretization by finite elements and an optimization loop for the discrete system. This solution algorithm is applied during the adaptive mesh refinement algorithm in each iteration.

Throughout this paper, we assume that the parameter identification problem described so far admits a (locally) unique solution. Moreover we assume the semi-linear form $a(\cdot,\cdot)(\cdot)$ and the observation operator C to be three times continuously differentiable and make the following assumption on the derivative $a'_u(\cdot,\cdot)(\cdot,\cdot)$:

Assumption 1 *In a neighborhood $B(u,q) \subset V \times Q$ of the solution (u,q) to problem (1,2) the derivative $a'_u(\cdot,\cdot)(\cdot,\cdot)$ is coercive, i.e. there exists a constant $\gamma > 0$ with*

$$(7) \qquad a'_u(v,p)(w,w) \geq \gamma \|w\|_V^2 \qquad \forall (v,p) \in B(u,q), \quad \forall w \in V.$$

We define an operator $A : V \times Q \to V'$ by:

$$(8) \qquad \langle A(u,q), \phi \rangle_{V' \times V} = a(u,q)(\phi) \quad \forall \phi \in V,$$

where $\langle \cdot, \cdot \rangle_{V' \times V}$ denotes the duality pairing between the Hilbert space V and its dual V'. The above assumption implies that the derivative $A'_u(u,q) : V \to V'$ is an isomorphism. Therefore, we obtain by the implicit function theorem in Banach spaces (see, e.g., Dieudonné [12]) the existence of a continuously differentiable solution operator S for the state equation in a neighborhood $Q_0 \subset Q$ of the solution to the problem (1,2). For all $q \in Q_0$ we have:

$$A(S(q), q) = f, \tag{9}$$

or equivalently

$$a(S(q), q)(\phi) = f(\phi) \quad \forall \phi \in V. \tag{10}$$

Using this solution operator S we define the reduced observation operator $c : Q_0 \to Z$ by:

$$c(q) := C(S(q)) \tag{11}$$

in order to reformulate the problem under consideration as an unconstrained optimization problem with the reduced cost functional $j : Q \to \mathbb{R}$:

$$\text{Minimize} \;\; j(q) := \tfrac{1}{2}\|c(q) - \bar{C}\|_Z^2 + \tfrac{\alpha}{2}\|q - \bar{q}\|_Q^2, \quad q \in Q. \tag{12}$$

Denoting by $G = c'(q)$ the Jacobian matrix of the reduced observation operator c, the first-order necessary condition for (12) reads:

$$G^* c(q) + \alpha q = G^* \bar{C} + \alpha \bar{q}. \tag{13}$$

In the following proposition we compute the Jacobian G.

Proposition 1 *Let the reduced observation operator c be defined as in (11). Then its partial derivatives can be computed as follows:*

$$\frac{\partial c_i}{\partial q_j}(q) = G_{ij} = C_i'(u)(w_j), \quad i = 1 \ldots n_m, \;\; j = 1 \ldots n_p, \tag{14}$$

with $u = S(q)$, C_i and c_i denote the components of the observation and the reduced observation operators respectively; G_{ij} denotes the entries of the Jacobian matrix $G = c'(q)$ and $w_j \in V$ is the solution to the following tangent problem:

$$a'_u(u,q)(w_j, \phi) = -a'_{q_j}(u,q)(1, \phi) \quad \forall \phi \in V. \tag{15}$$

Proof. Taking the derivatives of

$$a(S(q), q)(\phi) = f(\phi) \quad \forall \phi \in V, \tag{16}$$

we obtain:

$$a'_u(u, q)(S'_{q_j}(q)(1), \phi) + a'_{q_j}(u, q)(1, \phi) = 0 \quad \forall \phi \in V, \tag{17}$$

where $S'_{q_j}(q)(1)$ denotes the partial derivative of the solution operator S with respect to q_j.

By the chain rule we have:

$$G_{ij} = \frac{\partial c_i}{\partial q_j}(q) = C'_i(u)(S'_{q_j}(q)(1)). \tag{18}$$

To complete the proof we use the definition (15) of w_j. □

We will throughout suppose that the problem is non-degenerate in the following sense:

Assumption 2 *The Jacobian matrix G of the reduced observation operator c has full rank n_p in a neighborhood of the solution to problem (1,2).*

In the sequel we will also need the second derivative of the reduced cost functional. We have

$$\nabla^2 j(q) = G^* G + M, \tag{19}$$

where the matrix $M \in \mathbb{R}^{n_p \times n_p}$ is defined by

$$M := -\sum_{i=1}^{n_m} c''_i(q) R_i^{LS}. \tag{20}$$

Here, R_i^{LS} denotes the i-th component of the least-squares residual $R^{LS}(u)$ with $u = S(q)$.

We collect the necessary information for computation of M in the next proposition.

Proposition 2 *The entries M_{jk} of the matrix M defined in (20) can be computed by:*

$$\begin{aligned} M_{jk} = &-a''_{uu}(u, q)(w_j, w_k, z) - a''_{uq_j}(u, q)(w_k, 1, z) - a''_{q_j q_k}(u, q)(1, 1, z) \\ &-a''_{uq_k}(u, q)(w_j, 1, z) - \langle C''(u)(w_j, w_k), R^{LS}(u)\rangle_Z, \end{aligned}$$

where $u = S(q)$. Further, $w_j \in V$ is defined in (15) and $z \in V$ is the solution of the following adjoint equation:

$$a'_u(u, q)(\phi, z) = -\langle R^{LS}(u), C'(u)(\phi)\rangle_Z. \tag{21}$$

The proof of Proposition 2 is similar to the one of Proposition 1.

2.1 Discretization

We consider two- and three dimensional meshes consisting of *cells* K which are either triangles, tetrahedra, quadrilaterals, or hexahedra and constitute a non-overlapping covering of the computational domain:

$$\Omega = \bigcup K.$$

Note that this implies that the boundary $\partial\Omega$ of the domain is polygonal. The general case necessitates the treatment of curved cells and is neglected here.

The corresponding mesh is denoted by $\mathcal{T}_h$, where the mesh parameter h is defined as a cell-wise constant function by setting $h|_K = h_K$ and h_K is the diameter of K. The straight parts which make up the boundary ∂K of a cell K are called *faces*.

A mesh $\mathcal{T}_h$ is called regular, if it fulfills the standard conditions for shape-regular finite element mesh, see e.g. Ciarlet [9]. However, the cells are allowed to have nodes, which lie on midpoints of faces of neighboring cells. But at most one such *hanging node* is permitted for each face. On a regular mesh, we construct continuous finite element spaces V_h in the standard way, see e.g Ciarlet [9]. Only the case of hanging nodes requires some additional remark. There are no degrees of freedom corresponding to these irregular nodes and the value of the finite element function is determined by pointwise interpolation. This implies continuity and therefore global conformity. For implementation details see e.g. Carey & Oden [8].

For a given finite element space V_h the corresponding discrete solution $(u_h, q_h) \in V_h \times Q$ is determined by the constrained least squares problem (4,5). We assume the discrete analog of Assumption 1, in order to guarantee the existence of a continuously differentiable discrete solution operator S_h in a neighborhood $Q_{0,h} \subset Q$ of the solution to the discrete problem, i.e, there holds for all $q \in Q_{0,h}$:

$$a(S_h(q), q)(\phi_h) = f(\phi_h) \quad \forall \phi \in V_h. \tag{22}$$

As before, we turn the discrete problem (4,5) into an unconstrained minimization problem:

$$\text{Minimize} \quad j_h(q_h) := \frac{1}{2}\|c_h(q_h) - \bar{C}\|_Z^2 + \frac{\alpha}{2}\|q_h - \bar{q}\|_Q^2, \quad q_h \in Q. \tag{23}$$

We denote by $G_h = c_h'(q_h)$ the Jacobian of the discrete reduced function and assume again that it has full rank. The first-order necessary condition for (23) reads:

$$G_h^* c_h(q_h) + \alpha q_h = G_h^* \bar{C} + \alpha \bar{q}. \tag{24}$$

Therefore, the discrete solution (u_h, q_h) is determined by the system of equations constituted by the state equation (5) and the optimality condition (24).

Remark 2 The derivatives of the discrete observation operator c_h can be computed in the same way as in Proposition 1 and 2.

2.2 Optimization algorithm

Here, we describe the iterative solution of (24) on a fixed mesh $\mathcal{T}_h$. Let q_h^0 be an initial guess (which will be the solution on the previous mesh in the adaptive algorithm). Then we iterate

$$q_h^{k+1} = q_h^k + \delta q_h, \tag{25}$$

where δq_h is the solution of the linear problem

$$(G_h^* G_h + \alpha I + M_h)\delta q_h = G_h^*(\bar{C} - c_h(q_h^k)) + \alpha(\bar{q} - q_h^k). \tag{26}$$

Here I is the identity operator on Q, G_h is the Jacobian

$$G_h = c_h'(q_h^k), \tag{27}$$

and the definition of M_h leads to different variants of the optimization algorithm. The two typical choices of M_h are $M_h = 0$ or, as in Proposition 2, $M_h = -\sum_{i=1}^{n_m} c_h''(q_h^k)_i R_{h,i}^{LS}$ with the discrete analog R_h^{LS} of the least-squares residual R^{LS}. The first one leads to the Gauß-Newton algorithm which can be interpreted as the solution to the linearized minimization problem

$$\text{Minimize } \tfrac{1}{2}\|c_h(q_h^k) + G_h\delta q_h - \bar{C}\|^2 + \tfrac{\alpha}{2}\|q_h^k + \delta q_h - \bar{q}\|_Q^2. \tag{28}$$

The second choice leads to the full Newton algorithm and requires computation of the second derivatives in M_h, which may be obtained in the same way as in Proposition 2. For convergence theory of these algorithms see, e.g., [11, 14].

In the following we give the details of a typical iteration, focusing for simplicity on the Gauß-Newton method. We use a given tolerance ε in the stopping criterion.

Optimization Algorithm

1. Choose an initial parameter $q_h^0 \in Q$ and set $k = 0$
2. Compute $u_h^k \in V_h$, the solution of
$$a(u_h^k, q_h^k)(\phi_h) = f(\phi_h) \quad \forall \phi_h \in V_h$$
3. Compute $w_{j,h}^k \in V_h$, $j = 1 \dots n_p$, the solutions of
$$a'_u(u_h^k, q_h^k)(w_{j,h}^k, \phi_h) = -a'_{q_j}(u_h^k, q_h^k)(1, \phi_h) \forall \phi_h \in V_h$$
4. Compute the matrix G_h by:
$$(G_h)_{ij} = C'_i(u_h^k)(w_{j,h}^k)$$
5. Compute the residual $r_h \in Q$ as
$$r_h = G_h^*(\bar{C} - C(u_h^k)) + \alpha(\bar{q} - q_h^k)$$
6. If $\|r_h\|_Q \leq \varepsilon$ quit
7. Compute δq_h as solution of:
$$(G_h^* G_h + \alpha I)\delta q_h = r_h$$
8. Set $q_h^{k+1} = q_h^k + \delta q_h$
9. Increment k and go to 2.

Remark 3 In our practical realization, we use trust-region techniques in order to improve global convergence, see e.g. [7, 11, 14].

Remark 4 The cost for one step of the Gauß-Newton algorithm described above is proportional to the dimension of the parameter space Q. However, it is almost independent of the number of observations, in contrast to our former approach [6], where instead of n_p tangent problems (in step 3) n_m dual problems are solved in order to compute the matrix G_h.

3 A posteriori error estimation

3.1 An introductory example

We discuss the question of a posteriori error estimation for a simple linear example with one parameter and one observation, i.e. $Q = Z = \mathbb{R}$. The state equation, the observation operator and the dependency of the state variable on the parameter are all linear. The regularization parameter α is zero.

Our problem is formulated as follows: for a positive weight function $\omega \in L^2(\Omega)$ and a number $\bar{C}$ we minimize the least squares functional

$$\frac{1}{2}\Big|\int_\Omega \omega u \, dx - \bar{C}\Big|^2 \tag{29}$$

under the constraint

$$-\Delta u = qf \quad \text{in } \Omega = (0,1)^2, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{30}$$

with positive $f \in L^2(\Omega)$. The meaning of this example problem is, that we adjust the scaling of the load in such a way, that the weighted mean of the solution u attains a prescribed value.

Here, the solution operator is given by

$$u = S(q) = qu_1, \tag{31}$$

where u_1 is the solution of the state equation (30) for $q = 1$, i.e.

$$-\Delta u_1 = f \quad \text{in } \Omega, \qquad u_1 = 0 \quad \text{on } \partial\Omega. \tag{32}$$

Therefore the optimal parameter is simply given by:

$$q = \bar{C}\mu, \qquad \mu := \frac{1}{\int_\Omega \omega u_1 \, dx}. \tag{33}$$

By virtue of the maximum principle, μ is well defined. Similarly, the discrete optimal parameter is given by

$$q_h = \bar{C}\mu_h, \qquad \mu_h := \frac{1}{\int_\Omega \omega u_{1h} \, dx}, \tag{34}$$

where $u_{1h} \in V_h$ is the solution of the discrete analog of (32).

Now, the fundamental idea is to introduce the adjoint equation:

$$-\Delta y = -\mu\omega \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \partial\Omega. \tag{35}$$

Then we have, using the definitions of q, μ and (35):

$$\begin{aligned} q - q_h &= \bar{C}\mu - q_h\mu \int_\Omega \omega u_1 \, dx \\ &= \mu \int_\Omega \omega u_h \, dx + q_h(\nabla u_1, \nabla y) \\ &= -(\nabla u_h, \nabla y) + (q_h f, \nabla y). \end{aligned}$$

Therefore the error in parameter is related to the residual ρ of the discrete equation,

$$q - q_h = \rho(y), \qquad \rho(\phi) := (q_h f, \phi) - (\nabla u_h, \nabla \phi). \tag{36}$$

Before turning this result into a concrete a posteriori error estimate, we make some remarks.

Remark 5 A direct application of the approach in Becker, Rannacher & Kapp [5] would lead to consider the following adjoint equation:

$$-\Delta z = -(\bar{C} - \int_\Omega \omega u \, dx)\omega / z|_{\partial\Omega} = 0. \tag{37}$$

The right hand side in (37) is the derivative of the least squares functional. This implies $z = 0$ in our case, and therefore, the resulting error bound is useless.

Remark 6 The simple form and derivation of equality (36) are due to the structure of the problem: beside the linearity and $n_m = n_p = 1$, the fact that the observations are matched, i.e. that the least squares residual $C(u) - \bar{C}$ vanishes, is crucial. In the general setting later on, we use an approach based on a special Lagrange functional.

Remark 7 The quantity μ captures the stability of the problem, since $\mu = 1/c'(q)$ with the reduced functional c defined in (11). It can also be seen that $\mu = G^*(G^*G)^{-1}$, which is the correct expression in the general setting.

Proposition 3 *For the discretization of the simple example (29,30), we have the a posteriori error estimate:*

$$|q - q_h| \le \eta := \sum_{K \in \mathcal{T}_h} \rho_K \omega_K, \tag{38}$$

with the cell residual and cell weights defined by:

$$\rho_K = \|q_h f + \Delta u_h\|_K + \tfrac{1}{2} h_K^{-1/2} \|[\partial_n u_h]\|_{\partial K}, \tag{39}$$

$$\omega_K = \|y - i_h y\|_K + h_K^{1/2} \|y - i_h y\|_{\partial K}, \tag{40}$$

where the second term in (39) involves the jump of the normal derivative over the interior faces of the mesh and is understood to be zero on boundary faces. The weights are local interpolation errors involving an appropriate interpolation operator $i_h : V \to V_h$, see Clément [10].

Proof. We first remark, that by virtue of the Galerkin discretization,

$$(q_h f, \phi_h) - (\nabla u_h, \nabla \phi_h) = 0 \qquad \forall \phi_h \in V_h.$$

Using this and cell-wise integration by part, we obtain from (36) with an arbitrary $\phi_h \in V_h$:

$$\begin{aligned} q - q_h = \rho(y) = \rho(y - \phi_h) &= (q_h f, y - \phi_h) - (\nabla u_h, \nabla(y - \phi_h)) \\ &= \sum_K (q_h f + \Delta u_h, y - \phi_h)_K - \frac{1}{2}([\partial_n u_h], y - \phi_h)_{\partial K}. \end{aligned}$$

Using the Cauchy-Schwarz inequality on each cell and choosing $\phi_h = i_h y$, we obtain the announced result. □

Remark 8 For evaluation of this error estimator, the local interpolation error $y - i_h y$ has to be approximated. Later on, we use interpolation of the computed bilinear finite element solution y_h on the space of biquadratic finite elements on patches of cells.

Remark 9 It is well-known that residual estimators of the above form leads to overestimation already in the case of error control with respect to the energy-norm for the Poisson equation. A simple alternative for obtaining better error bounds is the direct evaluation of the error representation $q - q_h = \rho(y - i_h y)$ using some approximation of the interpolation error $y - i_h y$, see Remark 8.

3.2 A posteriori error estimation in the general setting

Here, we derive our a posteriori error estimator for the error in the parameters in the general setting. More precisely, we estimate the error with respect to a given error functional $E : Q \to \mathbb{R}$. The functional E is supposed to be three times continuously differentiable and we define its gradient by identification in the usual way:

$$\langle \nabla E(q), \delta q \rangle_Q = E'(q)(\delta q) \qquad \forall \delta q \in Q.$$

Our goal is to prove the following error representation:

$$E(q) - E(q_h) = \eta_h + R,$$

where η_h denotes the a posteriori error estimator to be developed and R is a cubic remainder term due to linearization. The precise result is given in Theorem 1. The evaluation of this error estimator require knowledge of certain second derivatives of the reduced observation operator c, which are computed in the last step of the optimization algorithm by using Newton's method. Further, we derive a simplified version of the error estimator, which is cheaper to evaluate and does not require the use of Newton's method; here additional remainder terms related to the least squares residual $R^{LS}(u) = \bar{C} - C(u)$ appear, see Theorem 2.

For the derivation of our a posteriori error estimator we make the following preparations:

In order to derive optimality conditions for the parameter identification problem (1,2) we introduce the Lagrangian $\mathcal{L}$:

$$\mathcal{L}(u, q, z) = \tfrac{1}{2}\|C(u) - \bar{C}\|^2 + \tfrac{\alpha}{2}\|q - \bar{q}\|_Q^2 + f(z) - a(u, q)(z), \tag{41}$$

for $u \in V$, $q \in Q$ and $z \in V$. The derivative of $\mathcal{L}$ is expressed by means of the residual functionals, defined, using the abbreviation $\xi := (u, q, z)$, by

$$\rho_u(\xi)(\phi) := f(\phi) - a(u, q)(\phi), \tag{42}$$

$$\rho_z(\xi)(\psi) := -\langle C'(u)(\psi), R^{LS}(u)\rangle_Z - a'_u(u, q)(\psi, z), \tag{43}$$

$$\rho_q(\xi)(\sigma) := \alpha\langle q - \bar{q}, \sigma\rangle_Q - a'_q(u, q)(\sigma, z), \tag{44}$$

where $\phi, \psi \in V$ and $\sigma \in Q$ are test functions. Now we have with $\delta\xi = (\delta u, \delta z, \delta q)$:

(45) $$\mathcal{L}'(\xi)(\delta\xi) = \rho_u(\xi)(\delta z) + \rho_z(\xi)(\delta u) + \rho_q(\xi)(\delta q).$$

For the error functional E, we introduce the following Lagrangian $\mathcal{M}$:

(46) $$\mathcal{M}(\xi, \chi) = E(q) + \mathcal{L}'(\xi)(\chi),$$

where we have introduced an additional set of variables $\chi = (v, p, y) \in V \times Q \times V$. As in Becker & Rannacher [4], we have an error representation expressed in the following proposition.

Proposition 4 *Let $x = (\xi, \chi) \in X = (V \times Q \times V)^2$ be a stationary point of $\mathcal{M}$, i.e.*

(47) $$\mathcal{M}'(x)(\delta x) = 0 \qquad \forall \delta x \in X.$$

Further let $X_h = (V_h \times Q \times V_h)^2 \subset X$ be a subspace and $x_h = (\xi_h, \chi_h) \in X_h$ the corresponding Galerkin solution

(48) $$\mathcal{M}'(x_h)(\delta x_h) = 0 \qquad \forall \delta x_h \in X_h.$$

Then, there holds

(49) $$E(q) - E(q_h) = \frac{1}{2}\mathcal{M}'(x_h)(x - \widehat{x}_h) + R,$$

where $\widehat{x}_h \in X_h$ is arbitrary and the remainder term R is given by:

(50) $$R = \frac{1}{2}\int_0^1 \mathcal{M}'''(x_h + s\, e)(e, e, e)s(s-1)ds,$$

with $e = x - x_h$.

Proof. We note, that ξ is a stationary point of $\mathcal{L}$, i.e.

(51) $$\mathcal{L}'(\xi)(\delta\xi) = 0 \qquad \forall \delta\xi \in V \times Q \times V$$

and ξ_h is the corresponding Galerkin solution

(52) $$\mathcal{L}'(\xi_h)(\delta\xi_h) = 0 \qquad \forall \delta\xi_h \in V_h \times Q \times V_h.$$

Therefore, we obtain:

(53) $$E(q) - E(q_h) = M(x) - M(x_h).$$

We rewrite the right hand side of (53) as follows:

$$\mathcal{M}(x) - \mathcal{M}(x_h) = \int_0^1 \mathcal{M}'(x_h + s\, e)(e)ds, \tag{54}$$

approximate the integral by the trapezoidal rule and obtain:

$$\mathcal{M}(x) - \mathcal{M}(x_h) = \frac{1}{2}\mathcal{M}'(x)(e) + \frac{1}{2}\mathcal{M}'(x_h)(e) + R, \tag{55}$$

where the remainder term R is defined in (50). The term $\mathcal{M}(x)(e)$ vanishes, and due to Galerkin orthogonality the term $\mathcal{M}(x_h)(e)$ can be replaced by $\mathcal{M}(x_h)(x - \widehat{x}_h)$ with $\widehat{x}_h$ arbitrarily chosen. This completes the proof. □

Next, we consider the derivative $\mathcal{M}'$. First, using (46) and the splitting of variables $x = (\xi, \chi)$, $\delta x = (\delta\xi, \delta\chi)$, we have:

$$\mathcal{M}'(x)(\delta x) = \mathcal{L}'(\xi)(\delta\chi) + E'(q)(\delta q) + \mathcal{L}''(\xi)(\delta\xi, \chi). \tag{56}$$

Therefore, the stationarity condition (47) splits into the following two systems of equations. First, the original variables ξ are determined by the stationarity of $\mathcal{L}$, which corresponds to variations $\delta\chi$. Second, the auxiliary variables χ are solution of the equations corresponding to variations $\delta\xi$. In order to rewrite the last terms, we introduce some additional residual functionals:

$$\rho_v(x)(\phi) := -a'_q(u, q)(p, \phi) - a'_u(u, q)(v, \phi), \tag{57}$$

$$\begin{aligned} \rho_y(x)(\psi) := \; & \langle C'(u)(v), C'(u)(\psi)\rangle_Z - \langle C''(u)(\psi, v), R^{LS}(u)\rangle_Z \\ & -a''_{uu}(u, q)(\psi, v, z) \\ & -a''_{uq}(u, q)(\psi, p, z) - a'_u(u, q)(\psi, y), \end{aligned} \tag{58}$$

$$\begin{aligned} \rho_p(x)(\sigma) := \; & E'(q)(\sigma) - a''_{uq}(u, q)(v, \sigma, z) - a''_{qq}(u, q)(\sigma, p, z) \\ & -a'_q(u, q)(\sigma, y) + \alpha\langle p, \sigma\rangle_Q. \end{aligned} \tag{59}$$

With these notations, we have:

$$\begin{aligned} \mathcal{M}'(x)(\delta x) = \; & \{\rho_u(\xi)(\delta y) + \rho_z(\xi)(\delta v) + \rho_q(\xi)(\delta p)\} \\ & +\{\rho_v(x)(\delta z) + \rho_y(x)(\delta u) + \rho_p(x)(\delta q)\}. \end{aligned} \tag{60}$$

At first glance, it seems as if application of Proposition 4 requires a huge coupled system to be solved. That this is not the case, is shown in Proposition 5. It turns out that the main afford in computation of the auxiliary solution (v, y, p) is already done in the optimization loop based on Newton's method as described in Subsection 2.2.

Proposition 5 *Let (u, z, q) be a stationary point of $\mathcal{L}$ (41) and let $H = (H_{jk})$ be the reduced Hessian defined by*

$$H = G^*G + \alpha I + M, \tag{61}$$

where $M = -\sum_{i=1}^{n_m} c_i''(q) R_i^{LS}$. Further, let $\{w_j\}_{1\le j\le n_p}$ be the tangent solutions defined in (15). Then, the auxiliary solution (v, y, p) is given by:

$$Hp = -\nabla E, \qquad v = \sum_{j=1}^{n_p} w_j p_j, \tag{62}$$

and y is the solution to the following equation:

$$\begin{aligned} a_u'(u,q)(\phi, y) &= \langle Gp, C'(u)(\phi)\rangle_Z - \langle C''(u)(\phi, v), R^{LS}(u)\rangle_Z \\ &\quad - a_{uu}''(u,q)(\phi, v, z) - a_{uq}''(u,q)(\phi, p, z) \quad \forall \phi \in V. \end{aligned} \tag{63}$$

Proof. (u, z, q) is a stationary point of $\mathcal{L}$ and consequently the residuals $\rho_u(x)$, $\rho_z(x)$ and $\rho_q(x)$ vanish. Due to the construction of v in (62) we obtain:

$$\begin{aligned} \rho_v(x)(\phi) &= -\sum_{j=1}^{n_p} p_j a_u'(u,q)(w_j, \phi) - a_q'(u,q)(p, \phi) \\ &= -\sum_{j=1}^{n_p} p_j \big(a_u'(u,q)(w_j, \phi) + a_{q_j}'(u,q)(1, \phi)\big). \end{aligned}$$

This sum vanishes because of the definition of w_j (15). Moreover, using Proposition 1 we obtain:

$$C_i'(u)(v) = \sum_{j=1}^{n_p} p_j C_i'(u)(w_j) = \sum_{j=1}^{n_p} G_{ij} p_j = (Gp)_i.$$

Therefore, $C'(u)(v) = Gp$ and $\rho_y(x)$ vanishes too due to the definition of y (63). Finally, in order to see, that $\rho_p(x) = 0$, we obtain using (62), (63) and the representations of G and M given by Proposition 1 and Proposition 2:

$$\begin{aligned} E_{q_j}'(q) &= -\sum_{k=1}^{n_m} G_{kj}(Gp)_k - \sum_{k=1}^{n_m} M_{jk} p_k - \alpha p_j \\ &= -\langle Gp, C'(u)(w_j)\rangle_Z + \langle C''(u)(w_j, v), R^{LS}(u)\rangle_Z \\ &\quad + a_{uu}''(u,q)(w_j, v, z) + a_{uq}''(u,q)(w_j, p, z) + a_{uq_j}''(u,q)(v, 1, z) \\ &\quad + a_{qq_j}''(u,q)(p, 1, z) - \alpha p_j \\ &= a_{uq_j}''(u,q)(v, 1, z) + a_{qq_j}''(u,q)(p, 1, z) - a_u'(u,q)(w_j, y) - \alpha p_j \\ &= a_{q_j}'(u,q)(1, y) + a_{uq_j}''(u,q)(v, 1, z) + a_{qq_j}''(u,q)(p, 1, z) - \alpha p_j. \end{aligned}$$

This completes the proof. □

Remark 10 The corresponding discrete set of auxiliary variables $\chi_h = (v_h, p_h, y_h) \in V_h \times Q \times V_h$ is constructed in the same way.

Therefore the main computational cost is the solution of the auxiliary problem (63). From Propositions 4 and 5 we obtain the following result using a suitable interpolation operator $i_h : V \to V_h$ (see, e.g., Clement [10]).

Theorem 1 *Let (u, z, q) be the solution of the parameter identification problem and let (v, y, p) be defined by equations (62) and (63). Further we denote by (u_h, z_h, q_h) and (v_h, y_h, p_h) the corresponding Galerkin solutions. Then, the following error representation holds:*

$$E(q) - E(q_h) = \frac{1}{2}\{\rho_u(\xi_h)(y - i_h y) + \rho_z(\xi_h)(v - i_h v)\} \tag{64}$$

$$+\frac{1}{2}\{\rho_v(x_h)(z - i_h z) + \rho_y(x_h)(u - i_h u)\} + R. \tag{65}$$

where R is the cubic remainder term defined in (50).

Proof. It is sufficient to remark that the residuals ρ_q and ρ_p vanish due to the finite dimension of Q. □

Theorem 1 gives a satisfactory result in the following sense: with the solution of only one auxiliary problem, we obtain an error representation with a cubic remainder term. For a discussion of the estimation of such remainder terms from linearization, see Vexler [15]. Moreover, no assumption on the smallness of the least squares residual $R^{LS}(u)$ is necessary. However, the Newton matrix H, and therfore computation of the adjoint solution z are required.

In the following, we provide a simpler error representation which can be used when the Gauß-Newton algorithm is employed in the optimization loop. For this purpose, we split the derivative $\mathcal{M}'$ into two parts:

$$\mathcal{M}'(x)(\delta x) = D(x)(\delta x) + T(x)(\delta x), \tag{66}$$

where the terms D and T are defined by

$$\begin{aligned} D(x)(\delta x) = &\{\rho_u(\xi)(\delta y) + \rho_z(\xi)(\delta v) + \rho_q(\xi)(\delta p)\} \\ &+\{\rho_v(x)(\delta z) + \tilde{\rho}_y(x)(\delta u) + \tilde{\rho}_p(x)(\delta q)\} \end{aligned} \tag{67}$$

and

$$T(x)(\delta x) = -\langle C''(u)(\delta u, v), R^{LS}(u)\rangle_Z - a''_{uu}(u, q)(\delta u, v, z)$$

$$-a''_{qu}(u, q)(\delta u, p, z) - a''_{uq}(u, q)(\delta q, v, z) \tag{68}$$

$$-a''_{qq}(u, q)(\delta q, p, z), \tag{69}$$

using the additional residuals:

$$\tag{70} \begin{aligned} \tilde{\rho}_y(x)(\psi) &:= \langle C'(u)(v), C'(u)(\psi)\rangle_Z - a'_u(u,q)(\psi,y), \\ \tilde{\rho}_p(x)(\sigma) &:= E'(q)(\sigma) - a'_q(u,q)(\sigma,y) + \alpha\langle p,\sigma\rangle_Q. \end{aligned}$$

The main idea is to replace the stationarity condition (47) by a simpler equation:

$$\tag{71} D(\tilde{x})(\delta x) = 0 \qquad \forall \delta x \in X.$$

Similar to Proposition 5 we obtain:

Proposition 6 *Let (u, z, q) be a stationary point of $\mathcal{L}$ (41) and let $\tilde{H} = (\tilde{H}_{jk})$ be the reduced Gauß-Newton matrix defined by*

$$\tag{72} \tilde{H} = G^*G + \alpha I.$$

Further, let $\{w_j\}_{1\le j\le n_p}$ be the tangent solutions defined in (15). Then, the solution $\tilde{x} = (u, q, z, \tilde{v}, \tilde{y}, \tilde{p}) \in X$ of (71) is given by:

$$\tag{73} \tilde{H}\tilde{p} = -\nabla E, \qquad \tilde{v} = \sum_{j=1}^{n_p} w_j \tilde{p}_j,$$

and $\tilde{y}$ is the solution to the following equation:

$$\tag{74} a'_u(u,q)(\phi,\tilde{y}) = \langle G\tilde{p}, C'(u)(\phi)\rangle_Z \qquad \forall \phi \in V.$$

The proof of Proposition 6 is similar to the one of Proposition 5.

The computation of $(\tilde{v}, \tilde{y}, \tilde{p})$ is cheaper in comparison with (v, y, p), because the matrix M is not required and the right hand side of (74) does not involve second derivatives of the semi-linear form a in contrast to (63). The corresponding error representation is formulated in the following theorem:

Theorem 2 *Let (u, z, q) be the solution of the parameter identification problem and let $(\tilde{v}, \tilde{y}, \tilde{p})$ be defined by equations (73) and (74). Further we denote by $\tilde{x}_h = (u_h, z_h, q_h, \tilde{v}_h, \tilde{y}_h, \tilde{p}_h)$ the corresponding Galerkin solution. Then, the following error representation holds:*

$$E(q) - E(q_h) = \frac{1}{2}\{\rho_u(\xi_h)(\tilde{y} - i_h\tilde{y}) + \tilde{\rho}_y(\tilde{x}_h)(u - i_h u)\} + \tilde{R} + P,$$

where $\tilde{R}$ is a remainder term defined by

$$\tag{75} \tilde{R} = \frac{1}{2}\int_0^1 D''(\tilde{x}_h + s\tilde{e})(\tilde{e},\tilde{e},\tilde{e})s(s-1)\,ds - \int_0^1 T'(\tilde{x}_h + s\tilde{e})(\tilde{e},\tilde{e})s\,ds,$$

with $\tilde{e} = \tilde{x} - \tilde{x}_h$. The additional remainder term P admits the estimate:

$$|P| \le \widetilde{C}\left(\|e_u\|_V + \|e_q\|_Q + \|\delta_h \tilde{v}\|_V + \|\delta_h \bar{z}\|_V\right)\|R^{LS}(u)\|_Z, \tag{76}$$

where $e_u := u - u_h$, $e_q := q - q_h$, $\delta_h \phi := \phi - i_h \phi$ is an interpolation error operator on V and $R^{LS}(u)$ is the least squares residual defined in (3). The normalized adjoint solution $\bar{z} \in V$ is determined by:

$$a'_u(u,q)(\phi,\bar{z}) = \langle -\tfrac{R^{LS}(u)}{\|R^{LS}(u)\|}, C'(u)(\phi)\rangle_Z \quad \forall \phi \in V, \tag{77}$$

if the least squares residual $R^{LS}(u)$ does not vanish; otherwise we set $\bar{z} = 0$. The constant $\widetilde{C}$ does not depend on the mesh parameter h nor on the measurements $\bar{C}$.

Proof. Similar to Proposition 4 we obtain the following error representation:

$$\begin{aligned} E(q) - E(q_h) &= \int_0^1 \mathcal{M}'(\tilde{x}_h + s\tilde{e})(\tilde{e})ds \\ &= \int_0^1 D(\tilde{x}_h + s\tilde{e})(\tilde{e})ds + \int_0^1 T(\tilde{x}_h + s\tilde{e})(\tilde{e})ds. \end{aligned}$$

We approximate the first integral by the trapezoidal rule, the second by the box rule and obtain

$$E(q) - E(q_h) = \frac{1}{2}D(\tilde{x}_h)(\tilde{e}) + \frac{1}{2}D(\tilde{x})(\tilde{e}) + T(\tilde{x})(\tilde{e}) + \tilde{R}, \tag{78}$$

with the corresponding remainder term $\tilde{R}$ defined in (75). The term $D(\tilde{x})(\tilde{e})$ vanishes, and due to Galerkin orthogonality the term $D(\tilde{x}_h)(\tilde{e})$ can be replaced by $D(\tilde{x}_h)(\tilde{x} - \widehat{x}_h)$ with $\widehat{x}_h$ arbitrarily chosen. We set

$$\widehat{x}_h = (i_h u, q, i_h z, i_h \tilde{v}, \tilde{p}, i_h \tilde{y}) \tag{79}$$

and obtain that the residuals ρ_q and $\tilde{\rho}_p$ vanish. We define the remainder term P by

$$P = T(\tilde{x})(\tilde{e}) + \tfrac{1}{2}\rho_z(\xi_h)(\delta_h z) + \tfrac{1}{2}\rho_v(\tilde{x}_h)(\delta_h \tilde{v}) \tag{80}$$

and it remains to prove the estimation (76). The term P has the following explicit form:

$$\begin{aligned} P = &-\langle C''(u)(e_u,\tilde{v}), R^{LS}(u)\rangle_Z - a''_{uu}(u,q)(e_u,\tilde{v},z) \\ &-a''_{qu}(u,q)(e_u,\tilde{p},z) - a''_{uq}(u,q)(e_q,\tilde{v},z) \\ &-a''_{qq}(u,q)(e_q,\tilde{p},z) - \langle C'(u)(\delta_h\tilde{v}), R^{LS}(u)\rangle_Z \\ &-a'_u(u,q)(\delta_h\tilde{v},z) - a'_u(u,q)(\tilde{v},\delta_h z) - a'_q(u,q)(\tilde{p},\delta_h z). \end{aligned} \tag{81}$$

We use the fact that $\bar{z} = z\,\|R^{LS}(u)\|_Z$ and rewrite (81) as

$$
\begin{aligned}
P = &-\langle C''(u)(e_u, \tilde{v}), R^{LS}(u)\rangle_Z - \|R^{LS}(u)\|_Z\, a''_{uu}(u,q)(e_u, \tilde{v}, \bar{z}) \\
&-\|R^{LS}(u)\|_Z\, a''_{qu}(u,q)(e_u, \tilde{p}, \bar{z}) - \|R^{LS}(u)\|_Z\, a''_{uq}(u,q)(e_q, \tilde{v}, \bar{z}) \\
&-\|R^{LS}(u)\|_Z\, a''_{qq}(u,q)(e_q, \tilde{p}, \bar{z}) + \langle C'(u)(\delta_h \tilde{v}), R^{LS}(u)\rangle_Z \\
&-\|R^{LS}(u)\|_Z\, a'_u(u,q)(\delta_h \tilde{v}, \bar{z}) - \|R^{LS}(u)\|_Z\, a'_u(u,q)(\tilde{v}, \delta_h \bar{z}) \\
&-\|R^{LS}(u)\|_Z\, a'_q(u,q)(\tilde{p}, \delta_h \bar{z}).
\end{aligned}
\tag{82}
$$

Using the Cauchy-Schwarz inequality and the continuity of the derivatives of C and a' completes the proof. □

Remark 11 For evaluation of the resulting error estimator, the local interpolation errors $u - i_h u$ and $\tilde{y} - i_h \tilde{y}$ are approximated by using higher order reconstructions, see Remark 8. For the localization of the error estimator, i.e. for the cell-wise representation of it, we use the same techniques as in Proposition 3.

Remark 12 In the case of perfect match, i.e. $R^{LS}(u) = 0$, the remainder term P vanishes and the error estimators resulting from Theorem 1 and from Theorem 2 are identical.

4 Numerical examples

In this section, we illustrate the usage of the a posteriori error estimator for some prototypical two-dimensional elliptic problems. Throughout, the discretization of the state equation uses piecewise bilinear finite elements on locally refined meshes consisting of quadrilaterals. The resulting nonlinear state equations are solved by Newton's method and the solution of the linear subproblems are computed using a standard multigrid algorithm. With these ingredients, the total numerical cost for solution on a given mesh behaves like $O(N)$, where N is the number of nodes. In our examples we also use point observations, yielding an unbounded observation operator C. This difficulty is overcome by using an appropriate regularization of the point functional, see [3] for details. All computation are done with the finite element library Gascoigne3D, see [2].

4.1 Example 1

We consider a convection-diffusion equation with unknown constant transport direction (q_1, q_2) in the unit square $\Omega = (0,1)^2$:

$$
\begin{aligned}
-\Delta u + q_1 u_x + q_2 u_y &= 2 \qquad \text{in } \Omega, \\
u &= 0 \qquad \text{on } \partial\Omega.
\end{aligned}
\tag{83}
$$

The parameters $(q_1, q_2) = (8, 8)$ are estimated using measurements given by the values of the state variable at five different points:

$$\begin{aligned} &\xi_1 = (0.25, 0.5), \quad \xi_2 = (0.5, 0.25), \\ &\xi_3 = (0.75, 0.5), \quad \xi_4 = (0.5, 0.75), \quad \xi_5 = (0.5, 0.5). \end{aligned} \tag{84}$$

The components of the corresponding observation operator C have the following form:

$$C_i(v) = v(\xi_i), \tag{85}$$

and the parameter identification problem is formulated as follows: For $(u, q) \in V \times Q$ with $V = H_0^1(\Omega)$ and $Q = \mathbb{R}^2$

$$\text{Minimize } \frac{1}{2} \sum_{i=1}^{5} (u(\xi_i) - \bar{C}_i)^2 \tag{86}$$

under the constraint (83), where $\bar{C}_i$ denote the components of the measurement vector $\bar{C} \in Z = \mathbb{R}^5$ and are given by the values of the state variable u for the exact parameter q, i.e. $\bar{C}_i = u(\xi_i)$.

Remark 13 Assumption 1 is fulfilled due to ellipticity of the state operator in (83). Moreover, the Jacobian matrix G has entries $G_{ij} = w_j(\xi_i)$, where the tangent solutions w_j are defined as in (15).

Due to the choice of the measurement points and to the fact, that these tangent solutions w_j are symmetric with respect to the line $y = x$, the matrix G has full rank $n_p = 2$. This proves that Assumption 2 is fulfilled, too.

Now we present numerical results using the a posteriori error estimator within the adaptive mesh refinement algorithm of Section 1. We estimate the error in the first component of $q \in Q$, i.e. we set:

$$E_1(q) := q_1. \tag{87}$$

First, we study the quality of the generated locally refined meshes. For comparison we also consider global refinement and refinement by an "energy" estimator for the state variable, which aims to control the error $u - u_h$ in $H^1(\Omega)$. Figure 1 compares the accuracy achieved on meshes resulting from these three types of refinement.

As seen from Figure 1, adaptive refinement based on the "energy"-estimator leads to almost the same reduction of the error as global refinement. However, the strategy based on our estimator leads to an obvious saving in the number of unknowns necessary to achieve a prescribed accuracy level.

Next, we investigate the quantitative behavior of the estimator which is important for error control. Comparison between the error in q_1 and the value

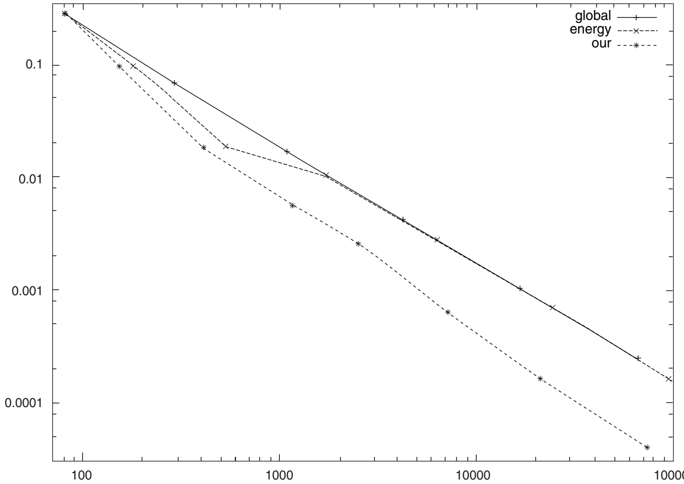

Fig. 1. Errors in q_1 for different refinement strategies vs. number of nodes

Table 1. Efficiency of the error estimator for the error functional $E_1(q) = q_1$

N	$E_1(q) - E_1(q_h)$	η	I_{eff}
81	2.80e-1	3.60e-1	0.78
153	9.59e-2	1.04e-1	0.92
407	1.82e-2	1.99e-2	0.91
1161	5.66e-3	5.39e-3	1.05
2487	2.47e-3	2.41e-3	1.02
7129	6.27e-4	6.14e-4	1.02
21091	1.63e-4	1.60e-4	1.02

of the estimator is shown in Table 1. There, the effectivity index of the error estimator is defined by

(88) $$I_{eff} := (E(q) - E(q_h))/\eta.$$

Computations are done on the same sequence of locally refined meshes as before (Figure 1).

Finally, we show some typical meshes resulting from application of the a posteriori error estimator for the error functionals $E_1(q) = q_1$, $E_2(q) = q_2$, and $E_3(q) = q_1^2 + q_2^2$ in Figure 2.

Remark 14 The presented example with the error functional $E_1(q) = q_1$ mimics a typical situation. One distinguishes between a set of parameters which are of primary interest, q_1, and others, which are also unknown and determined on the way. Depending on the choice of primary parameters, the meshes may differ significantly.

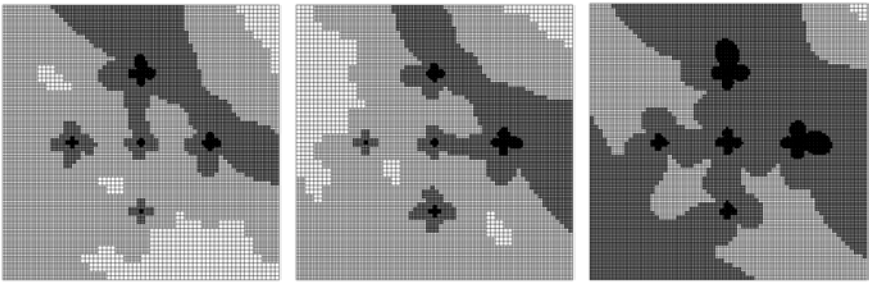

Fig. 2. *Typical meshes produced for $E_1(q) = q_1$ (left), $E_2(q) = q_2$ (middle) and $E_3(q) = q_1^2 + q_2^2$ (right)*

4.2 Example 2

We consider a diffusion-reaction equation with unknown constant coefficients $q = (q_1, q_2)$ on a cruciform domain Ω:

$$\begin{aligned} -q_1 \Delta u + q_2 u &= 2 \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{89}$$

The parameters (q_1, q_2) are estimated using measurements given by the point value in the middle of the domain Ω and the mean value of the state variable u. Therefor, the components of the corresponding observation operator C have the following form:

$$\begin{aligned} C_1(v) &= v(0,0), \\ C_2(v) &= \int_\Omega v\,dx. \end{aligned} \tag{90}$$

The parameter identification problem is formulated as follows: For $(u, q) \in V \times Q$ with $V = H_0^1(\Omega)$ and $Q = \mathbb{R}^2$

$$\text{Minimize } \ \tfrac{1}{2}(u(0,0) - \bar{C}_1)^2 + \tfrac{1}{2}\Big(\int_\Omega u\,dx - \bar{C}_2\Big)^2 \tag{91}$$

under the constraint (89), where $\bar{C}_i$ denote the components of the measurement vector $\bar{C} \in Z = \mathbb{R}^2$ and are given by the values of the observation operator C for the exact parameters $(q_1, q_2) = (1, 100)$, i.e. $\bar{C}_i = C_i(u)$. For these values of exact parameters, the solution u of (89) exhibits a visible boundary layer.

Typical meshes resulting for the two different error functionals $E_1(q) = q_1$ and $E_2(q) = q_2$ are shown in Figure 3.

As seen from Figure 3, the resulting mesh refinement is quite different for the two different error functionals. In both cases there is a strong refinement along the boundary; however, the mesh close to the point of evaluation for the functional C_1 is only refined in case of E_2. This behavior seems to be

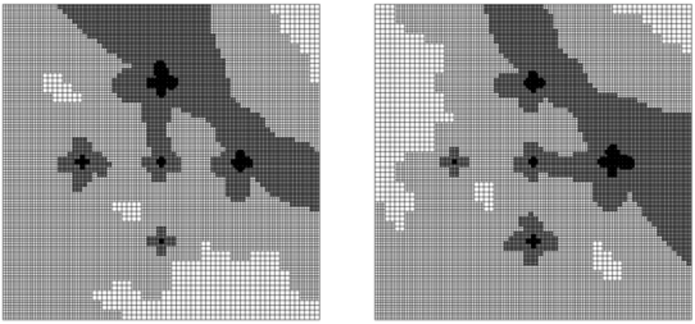

Fig. 3. *Typical meshes produced for $E_1(q) = q_1$ (left) and $E_2(q) = q_2$ (right)*

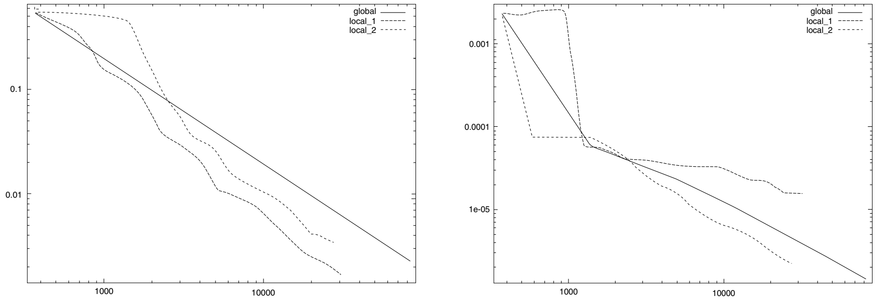

Fig. 4. *Errors in q_1 (left) and in q_2 (right) for different refinement strategies vs. number of unknowns*

surprising at the first glance. However, the missing refinement close to the point in the middle for E_1 is explained by the fact, that the value of C_1 is almost entirely determined by q_2 due to the boundary layer character.

In Figure 4 a comparison of the error in q_1 and the error in q_2 for the three strategies global mesh refinement, mesh refinement according to E_1, and mesh refinement according E_2 is made.

Obviously, for each error functional the corresponding refinement strategy is the most efficient one. In the second case (error in q_2) the mesh refinement according to E_1 is even worse than global refinement. This is in accordance with the explanation for the different refinement of the meshes shown in Figure 3.

4.3 Example 3

In our last example, we investigate a simple case of non-vanishing least squares residual. Here, we consider a convection-diffusion-reaction equation with a nonlinear reaction term:

$$(92) \qquad \begin{aligned} -\Delta u + e^{q_1 u} + q_2 u_x + q_3 u_y &= 2 \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned}$$

Again, we choose $\Omega = (0, 1)^2$. The parameters $(q_1, q_2, q_3) \in Q = \mathbb{R}^3$ are estimated using measurements given by the values of the state variable at the same points ξ_i (84) as in Example 1. Therefore, the components of the corresponding observation operator C have the form:

$$(93) \qquad C_i(v) = v(\xi_i).$$

The parameter identification problem is formulated as follows: For $(u, q) \in V \times Q$ with $V = H_0^1(\Omega)$ and $Q = \mathbb{R}^3$

$$(94) \qquad \text{Minimize } \tfrac{1}{2} \sum_{i=1}^{5} (u(\xi_i) - \bar{C}_i)^2 + \tfrac{\alpha}{2} \|q - \bar{q}\|_Q^2$$

under the constraint (92), where $\bar{C}_i$ denote the components of the measurement vector $\bar{C} \in Z = \mathbb{R}^5$ and $\bar{q} \in Q$ is a given reference parameter. In contrast to Example 1 and 2 we consider the case of $\alpha > 0$. The data ($\bar{C}$, $\bar{q}$ and α) are chosen, such that the least squares residual $R^{LS}(u)$ does not vanish.

Table 2. Comparison between the error in $E(q)$ and the values of the estimators η_1 and η_2 for the data with $\|R^{LS}(u)\|_Z \approx 3 \cdot 10^{-4}$

N	$E(q) - E(q_h)$	η_1	I_{eff}^1	η_2	I_{eff}^2
25	3.27e-1	6.92e-2	4.72	4.87e-2	6.73
81	7.53e-2	8.53e-2	0.88	7.80e-2	0.96
289	1.84e-2	1.87e-2	0.98	1.74e-2	1.06
1089	4.57e-3	4.51e-3	1.01	4.21e-3	1.09
4225	1.14e-3	1.12e-3	1.02	1.04e-3	1.10

Table 3. Comparison between the error in $E(q)$ and the values of the estimators η_1 and η_2 for the data with $\|R^{LS}(u)\|_Z \approx 4 \cdot 10^{-2}$

N	$E(q) - E(q_h)$	η_1	I_{eff}^1	η_2	I_{eff}^2
25	1.87e-1	2.06e-1	0.90	5.31e-2	3.52
81	3.58e-2	4.74e-2	0.76	1.26e-2	2.84
289	8.24e-3	8.71e-3	0.94	3.29e-3	2.50
1089	1.99e-3	2.00e-3	0.99	8.31e-4	2.39
4225	4.70e-4	4.85e-3	0.97	2.08e-4	2.26

Here, the a posteriori error estimators η_1 and η_2 based on the error representation from Theorem 1 and Theorem 2 respectively, differ. In order to illustrate the difference between the estimators, we consider two cases with different magnitude of $\|R^{LS}(u)\|_Z$. In both cases the error functional is $E(q) = q_2$. The values of the estimators η_1 and η_2 are shown in Table 3 and Table 2. The corresponding effectivity indexes I_{eff}^1 and I_{eff}^2 are defined as in (88).

References

1. Becker, R.: Adaptive Finite Elements for Optimal Control Problems. Habilitationsschrift, Institut für Angewandte Mathematik. Universität Heidelberg, 2001
2. Becker, R., Braack, M.: Gascoigne3D, A finite element toolkit for flow problems. SFB-preprint, Heidelberg, 2003
3. Becker, R., Rannacher, R.: A feed-back approach to error control in finite element methods: Basic analysis and examples. [J] East-West J. Numer. Math. **4**(4), 237–264 (1996)
4. Becker, R., Rannacher, R.: An optimal control approach to a posteriori error estimation in finite element methods. In: Acta Numerica 2001 A. Iserles, (ed), Cambridge: Cambridge University Press, 2001
5. Becker, R., Kapp, H., Rannacher, R.: Adaptive finite element methods for optimal control of partial differential equations: Basic concept. SIAM J. Cont. Opt. **39**(1), 113–132 (2000)
6. Becker, R., Vexler., B.: Mesh Adaptation for Parameter Identification Problems. In: Proc. of the 8th Conference on Control of Distributed Parameter Systems Graz, July 15-21, 2001, Intern. Series of Numerical Mathematics, **143**, 41–56 (2002)
7. Bock, H.G.: Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen. Bonner Mathematische Schriften, 1987
8. Carey, C.F., Oden, J.T.: Finite Elements, Computational Aspects. Vol. III. Prentice-Hall, 1984
9. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Amsterdam: North-Holland Publishing Company, 1978
10. Clement, Ph.: Approximation by finite element functions using local regularization. [J] Revue Franc. Automat. Inform. Rech. Operat. **9**(R-2), 77–84 (1975)
11. Dennis, J.E., Schnabel, R.B.: Numerical methods for unconstrained optimization and nonlinear equations. Number 16 in Classics in Applied Mathematics. SIAM, Society for Industrial and Applied Mathematics, 1996
12. Dieudonné, J.: Fondation of Modern Analysis. New York: Academic Press, 1960
13. Liu, W., Yan, N.: A posteriori error estimates for convex boundary control problems. SIAM J. Numer. Anal. **39**(1), 73–99 (2001)
14. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer Series in Operations Research, New York: Springer, 1999
15. Vexler, B.: A posteriori Fehlerschätzung und Gitteradaption bei Finite-Elemente-Approximationen nichtlinearer elliptischer Differentialgleichungen. Diplomarbeit, Universität Heidelberg, 2000